Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 38 (2012) 1599 – 1605

Procedia Engineering

www.elsevier.com/locate/procedia

International Conference on Modeling Optimisation and Computing (ICMOC-2012)

Temporal Based Approach for Face Information: A Modified Solution to AAM

C.Jaiseeli,K.Naga Prasad,S.Kiran Kumar,Vootla M Manoj,N.R.Raajan a*,M.V.Priya

Department of Electronics & Communication Engineering
School of Electrical & Electronics Engineering
SASTRA University
Thanjavur, Tamilnadu, India

Abstract

To gain High efficiency, Accurate alignment and for Handling Face Deformations we use Temporal Matching Active Appearance Model (TM-AAM) . The major component in judging a person's intention, personality and effective state are facial expressions and human face is a rich source of information for the viewer. The Specialty of TM-AAM is that it can generalize to any valid example as it contains a statistical model of the shape and grey level appearance of the object of interest. Robust and fast deformable image matching can be done by TM-AAM by combining texture and shape models. Image matching can be done by measuring the residuals and use the model for current parameter changes prediction to yield a better fit. This technique is widely used in high surveillance cameras, fugitive identification, security checks and for medical interpretations.

eywords: AAM, detection, face, model, shape parameters, temporal matching.

*Corresponding author. Tel.:+91-994-080-8955.

E-mail address: nrraajan@ece.sastra.edu. nagaprasad75@gmail.com.kirankumarsrr@gmail.com. manomurali91@gmail.com

1. Introduction

Interpretation of images of faces using model based approaches are now becoming increasingly attractive. Robust results can be obtained by using these approaches by constraining solutions to be valid instances of the model. In order to obtain these results the model of the face appearance should be as complete as possible i.e... the model should be able to synthesize a close approximation of the target face.

Many modelling approaches are present to interpret objects but few of them use full, photo-realistic models which are matched directly by minimizing difference between image under interpretation and the one synthesized by the model. Although suitable photorealistic models exist they typically involve a large set of parameters (50-100) in order to deal with the variability of the target objects. Optimization using such methods with high dimensional space is realistic but considerably slow.

In this paper we have used a technique called Temporal matching Active Appearance Model where direct optimization approach is used which is robust ,rapid and accurate. In this method solving of a general optimization problem each time a new image is given is not done. Instead of this the fact that the optimization problem is same each time is taken into account and these similarities are learnt offline. This approach is rather statistical than physical.

In this paper the idea of image interpretation through synthesis and previous related work is described. To achieve this we have to build compact models of face appearance which should be capable of generating synthetic examples similar to those of training set. This method can be used in variety of applications, here we have taken example of face images. Active Appearance Model algorithm is described in detail and its performance is demonstrated.

1.1. History

Many model based approaches have been described to the interpretation of images of deformable objects in the recent years. The basis for a broad range of applications is provided by the model by explaining the appearance of a given image in terms of a compact set of model parameters. The efficient method to interpret an image is to find best match between image and a model. The most common method is to allow a prototype to vary according to some physical model. The volume model is proposed by Bajscy and kovacic [4] which deforms elastically to generate new examples.

Principal component analysis is used to describe the face images in terms of a set of basic functions or 'Eigen faces' which was proposed by Turk and Pentland [5] . Using correlation methods the model can be matched to an image easily but the Eigen Faces is not robust to changes in shape and also to pose and expression. Poggio [6] and Co-workers made to fit the model to an unseen view by a stochastic optimisation procedure which is robust because of the quality of the synthesised images but is slow. A 3D model of the grey-level was described by Cootes [7] for full synthesis of shape and appearance. Nastar [8] also described a related model of the 3D grey-level surface by combining both physical and statistical modes of variation.

Cootes used Active shape model (ASM) to model shape and local grey-level appearances to locate flexible objects in new images. Edwards [3] worked to produce a combined model of shape and grey-level appearance which is the further extension of ASM. Thus AAM algorithm came into existence with the extension of Eigen feature model and Active shape model to overcome the defects in it. Sclaroff and Isidoro [9] have demonstrated 'Active Blobs' which uses only a single example rather than training set of examples. This approach is useful for 'bootstrapping'. The Temporal matching Active appearance model (AAM) overcomes the disadvantage of ASM which uses only shape constraints by taking all the available texture information across the target object. AAM has the ability to attain high accuracy and efficiency for handling face deformations than ASM for face tracking.

2. Methodology

In this paper we describe a technique for Face tracking and recognition called as Active Appearance model.. The main concept of this algorithm is that it uses the difference between the current estimate of appearance and the target image. Here we show how these appearance models are generated. Shape variations are combined with appearance variations in a shape – normalized frame to generate these models [1] .

Step 1: We create a training set of labelled images by marking key landmark points on each face. The shape of the object is determined by the labelled points.
Step 2: We create a statistical model of shape variation . All the sets are aligned on common co-ordinate frame and each are represented by a vector , x.
Step 3: Principal component analysis (PCA) is applied to the above data.

Now any example can be approximated using:

$$x = \bar{x} + P_s b_s \tag{1}$$

$\bar{x}$ = mean shape
P_s = set of orthogonal modes of variation
b_s = set of shape parameters

Step 4: Now Grey-level appearance statistical model is created by warping each example image so that its control points match the mean shape . Grey level information is sampled from the shape normalized image over the region covered by the mean shape.
Step 5: The effect of global lighting variation is minimized by normalizing the example samples by a scaling factor α, and offset β,

$$g = (g_{im} - \beta.1) / \alpha \tag{2}$$

α and β are chosen such that the vector is matched to the normalized mean . If g is the mean of the normalized data then the values of α and β are given by

$$\alpha = g_{im}\,\bar{g}\,, \quad \beta = (g_{im}.1) / n \tag{3}$$

where n is the number of elements in the vectors.
Step 6: A solution can be obtained by using one of the examples as the first estimate of the mean ,Aligning others to it and re-estimating the mean and iterating. By applying PCA to this normalized data we obtain the following:

$$g = \bar{g} + P_g b_g \tag{4}$$

Where $\bar{g}$ = normalized gray level vector
P_g = set of orthogonal modes of variation
b_g = set of gray-level parameters

Step 7: Now shape and appearance of any example can be given by using these vectors bg and bs. Due to the correlation between the shape and grey level variations a further PCA (principal component analysis is applied to the data . Thus a concatenated vector is formed as:

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \bar{x}) \\ P_g^T (g - \bar{g}) \end{pmatrix} \quad (5)$$

Where W_s is the diagonal matrix of weights for each shape parameter.
We apply a further PCA to these vectors to obtain

$$b = Qc \quad (6)$$

Q = eigen vectors
c = appearance parameters controlling both shape and gray-levels of the model.
Since the model is linear we can express shape and gray levels directly as functions of c

$$x = \bar{x} + P_s Q_s W_s c \quad , \quad g = \bar{g} + P_g Q_g c \quad (7)$$

where

$$Q = \begin{pmatrix} Q_s \\ Q_g \end{pmatrix} \quad (8)$$

Now an example image can be formed by generating the shape free gray level image from the Vector g and warping around the control points given by x. Each element of is displaced from its optimum value on each training example and each image with displaced shape is sampled. The weight is given by the RMS change in per unit change in shape parameter .

When a new image is given with a set of landmark points , an approximation can be generated by using the model. The above steps are used to get b by combining shape and gray level parameters.

As Q is orthogonal the combined appearance model parameters c are given by,

$$c = Q^T b \quad (9)$$

The full reconstruction is done by applying (7) and inverting gray-level normalization, applying pose to the points and projecting gray level vector onto the image.

2.1.Active Appearance Model using Temporal Matching

Now when an image is given for interpreting and appearance model as above is generated ,the scheme adjusts the model parameters so that a synthetic example is generated ,which matches closely to the new image. Here interpretation is treated as optimization problem in which difference between new image and one synthesized by the appearance model. Difference vector is given by:

$$\delta I = I_i - I_m \quad (10)$$

$I_i =$ vector of grey -level values of the image
$I_m =$ vector of grey-level values of the current model parameters

To get the best match between image and model ,we have to minimize the magnitude of the difference vector , $\Delta = |\delta I|^2$, by changing the model parameters c. The spatial pattern in δI ,has information about how the parameters should be changed in order to get good fit. So a simple relation between δI and δc is taken which is linear,

$$\delta c = A\delta I \quad (11)$$

To get A we do multiple multivariate regression on sample of known model displacements δc and corresponding difference images δI. Displacement in 2D positions, scale and orientation are also modeled. Difference is calculated as: let $c0$ be the known appearance model parameters for the current image. When parameters are displaced by a known amount δc new parameters are $c = c0 + \delta c$,for these parameters shape x and normalized gray levels g_m are generated. We sample from image warping at points x and forming normalized sample g_s. The sample error is given by $\delta g = g_s - g_m$.
We obtain a relation :

$$\delta c = A\delta g \quad (12)$$

The procedure of iterative model refinement is given as:

- Calculate error vector $\delta g0 = g_s - g_m$
- Calculate current vector $E0 = |\delta g0|^2$
- Compute predicted displacement $\delta c = A\delta g0$
- Set k = 1
- Assume $c1 = c0 - Q\delta c$
- Sample image at this new prediction, and calculate $\delta g1$
- If $|\delta g1|^2 < E0$,accept new estimate c1

This method is repeated till no improvement is made to error $|\delta g|^2$.
Inter-frame local appearance between frames is done by temporal matching [2]. It infuses good generalizability and does not suffer from mis-matched points. In the previous frame we select some feature points with salient local appearances and match it with the current frame.The tracker becomes resistant to global illumination changes as we match the local patches.

$$E = E_a + W_t E_t \quad (13)$$

W_t =controls strength of temporal matching constraint
E_a = AAM cost function
E_t =Temporal matching cost function.
E gives resultant function of AAM with temporal matching.

3. Experimental Results

Here we have tracked our face in a real time video using MATLAB integrated with webcam as shown in Fig 1.

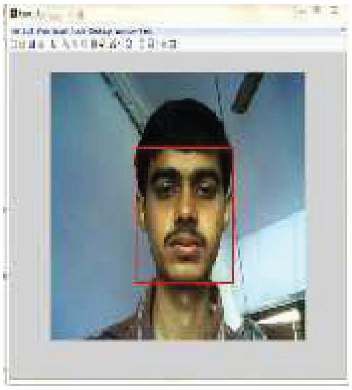

Fig 1

We created a training set for faces of three images each labelled with 17 points around the main features. Fig 2 , Fig 3 , Fig 4 are the input images.

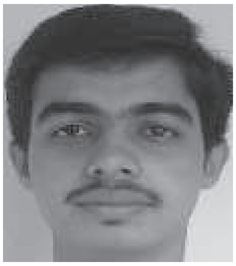

Fig 2

Fig 3

Fig 4

The table shown below gives Database of the (x, y) coordinates of 17 featured points of above images.

Fig 5(a)		Fig 5(b)		Fig 5(c)	
X1	Y1	X2	Y2	X3	Y3
115	223	100	191	94	182
243	222	227	190	230	182
90	253	76	223	68	227
148	259	132	223	125	222
205	258	190	225	193	226
265	253	247	224	250	227
56	316	35	281	38	293
142	317	130	276	120	288
178	317	163	274	152	284
213	314	194	275	192	284
289	314	276	278	286	289
128	370	118	332	105	347
177	352	163	319	157	324
177	371	161	333	157	347
177	390	161	353	159	371
223	367	204	336	211	346
177	447	162	420	161	430

By reading the values from the database landmark points are labelled at the respective feature points on the faces using MATLAB as shown below in Fig 5(a) ,Fig 5(b) , Fig 5(c).

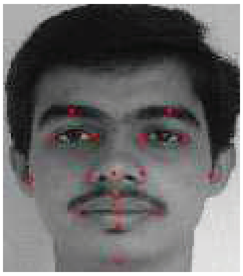

Fig 5(a)

Fig 5(b)

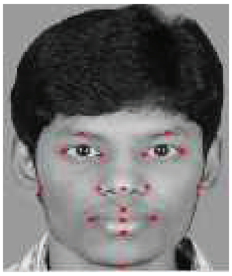

Fig 5(c)

By comparing the pixel values of the feature points of the database image and test image the person is identified.

4. Conclusion

Face Tracking based on Temporal matching Active appearance model (TM-AAM) has advantages of accurate alignment , high efficiency and effectiveness of handling face deformations. By using the vector coordinates approach we match the feature points on the face from previous frame onto the present frame. By using AAM algorithm we can overcome the defects of varying illumination and cluttered backgrounds. Thus AAM algorithm is simpler and more efficient than other techniques.

References

[1] T.F.Cootes, G.J.Edwards, C.J.Taylor, "Active Appearance Models", *computer vision and image Understanding.*

[2] M.Zhou, L.Liang, J.Sun and Y.Wang,"AAM Based Face Tracking with Temporal Matching and Face Segmentation",work.

[3] G.J.Edwards,C.J.Taylor and T.Cootes. "Learning to identify and track faces in image sequences",Ina 8th *British Machine Vision Conference,* Colchester,U.K,1997.

[4] Bajcsy and A.Kovacic. Multiresolution elastic matching. *Computer graphics and image processing,* 46:1-21,1989.

[5] M. Turk and A. Pentland . Eigenfaces for recognition. *Journal of cognitive neuroscience,* 3(1):71-86,1991.

[6] M.J. Jones and T. Poggio . Multidimensional Morphable models . In 6th *International conference on Computer vision,*pages 683-688,1998.

[7] T. Cootes and C. Taylor. Modelling object appearance using grey-level surface. In E.Hancock ,Editor 5th *British machine vision conference,*pages 479-488 , York , England , September 1994.

[8] C . Nastar , B. Moghaddam and A.Pentland. Generalized image matching :"Statistical learning of physically- based deformations" . In 4th *European conference on computer vision ,* volume 1 ,pages 589-598,CambridgeU.K , 1996.

[9] S. Sclaroff and J. "Isidoro.Active blobs". In 6th *International conference on computer vision.*pages 1146-53, 1998.